Monotonicity in the Sample Size of the Length of Classical Confidence Intervals

Abram M. Kagan[a] and Yaakov Malinovsky[b,*]

[a] Department of Mathematics, University of Maryland, College Park, MD 20742, USA

[b] Department of Mathematics and Statistics , University of Maryland Baltimore County, Baltimore, MD 21250, USA

[a] *email*: amk@math.umd.edu

[b] *email*: yaakovm@umbc.edu, *Corresponding author

Summary

It is proved that the average length of standard confidence intervals for parameters of gamma and normal distributions monotonically decrease with the sample size. The proofs are based on fine properties of the classical gamma function.

Key words: Gamma function; Location and scale parameters; Stochastic monotonicity

1 Introduction and Lemmas

In recent issues of the Bulletin of the IMS (see Shi (2008), DasGupta (2008)), a discussion was held on the behavior of standard estimators of parameters as functions of the sample size n. If $R(\widetilde{\theta}_n, \theta)$ is the risk of an estimator $\widetilde{\theta}_n$ constructed from a sample of size n, a very desirable property of $\widetilde{\theta}_n$ would be

$$R(\widetilde{\theta}_{n+1}, \theta) \leq R(\widetilde{\theta}_n, \theta) \tag{1}$$

for all θ. Unfortunately, even when (1) holds, it is very difficult to prove it for more or less general class of estimators and/or families.

One of few examples of classical estimators with a monotonically (in n) decreasing risk is the Pitman estimator of a location parameter. Let $(X_1, \ldots, X_n)$ be a sample from population $F(x-\theta)$ and let $t_n = t_n\,(X_1, \ldots, X_n)$ be the Pitman estimator corresponding to an (invariant) loss function $L(\widetilde{\theta}, \theta) = L(\widetilde{\theta} - \theta)$. The corresponding risk $R(\widetilde{\theta}_n)$ of any equivariant estimator $\widetilde{\theta}_n$ is constant in θ and by the very definition of t_n, $R(t_n) \geq R(t_{n+1})$ for any F. More deeper result holds for the Pitman estimator corresponding to the quadratic loss $L(\widetilde{\theta}-\theta) = (\widetilde{\theta}-\theta)^2$. If $\displaystyle\int x^2 dF(x) < \infty$, then for any n, $Var(t_n) < \infty$ and

$$nVar(t_n) \geq (n+1)Var(t_{n+1}). \tag{2}$$

The proof of (2) in Kagan *et al.* (2011) is based on a lemma of general interest from Artstein *et al.* (2004). The inequality was used in studying a geometric property of the sample mean in Kagan and Yu (2009).

Turning to the interval estimation of parameters, one finds a very natural loss function, namely the length of a confidence interval. Here we study the risk, i. e., the average length of the standard confidence intervals for the scale parameter β of a gamma distribution $Gamma\,(a, \beta)$ and for the mean μ and variance σ^2 of a normal distribution $N(\mu, \sigma^2)$. Though our results are new, to the best of our knowledge, their interest is more methodological than applied. Notice, however, that the distributions we study are often used in different applications.

It is proved that the average length of the standard confidence interval of a given level $1 - \alpha$ monotonically decreases with the sample size n. Though the monotonicity seems a very natural property, the proofs are based on fine properties of the gamma function and are nontrivial.

We write $X \sim Gamma\,(a, \beta)$ if the probability density function of X is

$$f\,(x; a, \beta) = \frac{1}{\beta^a \Gamma(a)} x^{a-1} e^{-x/\beta}, \;\; x > 0, a > 0, \beta > 0. \tag{3}$$

Lemma 1. *Let $X_1 \sim Gamma\,(a_1, \beta_1),\ X_2 \sim Gamma\,(a_2, \beta_2)$. If $a_1 < a_2, \beta_1 > \beta_2$, there exists a unique $x^* = x^*(a_1, a_2, \beta_1, \beta_2)$ such that the distribution functions F_1 of X_1, and F_2 of X_2 have the following properties:*

$$F_1(x) > F_2(x) \quad for \;\; x < x^* \;\; and \;\; F_1(x) < F_2(x) \quad for \;\; x > x^*. \tag{4}$$

Proof. Plainly, if $a_1 < a_2, \beta_1 > \beta_2$, then

$$\lim_{x \to 0} \frac{f(x; a_1, \beta_1)}{f(x; a_2, \beta_2)} = \lim_{x \to \infty} \frac{f(x; a_1, \beta_1)}{f(x; a_2, \beta_2)} = \infty. \tag{5}$$

Thus, for sufficiently small x, say $x < x_1$, $F_1(x) > F_2(x)$ and for for sufficiently large x, say $x > x_2$, one has $1 - F_1(x) > 1 - F_2(x)$, so that $F_1(x) < F_2(x)$. Therefore, there is a point x^* with $F_1(x^*) = F_2(x^*)$. We shall show that such an x^* is unique.

Take now the ratio $\dfrac{f(x; a_1, \beta_1)}{f(x; a_2, \beta_2)} \propto x^{a_1 - a_2} e^{-x(1/\beta_1 - 1/\beta_2)}$. Its log-derivative (with respect to x), $\dfrac{a_1 - a_2}{x} - \left(\dfrac{1}{\beta_1} - \dfrac{1}{\beta_2}\right)$ vanishes at the only point $x_0 = \dfrac{a_1 - a_2}{1/\beta_1 - 1/\beta_2}$. Let x', x'' with $x' < x''$ be two points with $f(x'; a_1, \beta_1) = f(x'; a_2, \beta_2)$ and $f(x''; a_1, \beta_1) = f(x''; a_2, \beta_2)$. Plainly, $x' < x_0 < x''$ and for all $x \in (x', x'')$ one has

$$\frac{f(x; a_1, \beta_1)}{f(x; a_2, \beta_2)} < 1. \tag{6}$$

Since for $x < x'$, $F_1(x) > F_2(x)$ and for $x > x''$, $1 - F_1(x) > 1 - F_2(x)$, any x with $F_1(x) = F_2(x)$ must satisfy $x' < x < x''$. If there are two points x_1^* and x_2^* with $F_1(x_i^*) = F_2(x_i^*)$, $i = 1, 2$, then by virtue of the Rolle theorem, there is a point $\widetilde{x}_0 \in (x_1^*, x_2^*)$ with $F_1'(\widetilde{x}_0) - F_2'(\widetilde{x}_0) = f(\widetilde{x}_0; a_1, \beta_1) - f(\widetilde{x}_0; a_2, \beta_2) = 0$, i.e., $\dfrac{f(\widetilde{x}_0; a_1, \beta_1)}{f(\widetilde{x}_0; a_2, \beta_2)} = 1$. However, for all $x \in (x', x'')$, the inequality (6) holds. The contradiction proves the uniqueness of x^* with $F_1(x^*) = F_2(x^*)$. □

For a special case of semi-integers a_1, a_2 (i. e., for the chi-squared distribution) the result of Lemma 1 was obtained in Székely and Bakirov (2003) by different arguments.

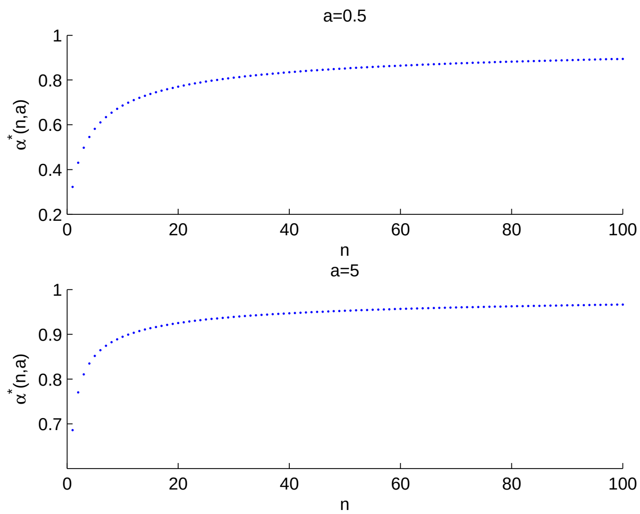

Figure 1: The critical values of α^* as a function of n

Since $F_1(x^*) = F_2(x^*) = \alpha^* = \alpha^*(a_1, a_2, \beta_1, \beta_2)$, say, the following is a straightforward Corollary of Lemma 1.

Corollary 1. *If $\gamma_{a_i,\beta_i;\alpha}$ is the quantile of order α of Gamma (a_i, β_i), $i = 1, 2$, then for $\alpha < \alpha^*$, $\gamma_{a_1,\beta_1;\alpha} < \gamma_{a_2,\beta_2;\alpha}$ and $\gamma_{a_1,\beta_1;1-\alpha} > \gamma_{a_2,\beta_2;1-\alpha}$.*

The next Lemma deals with a fine property of the classical gamma function.

Lemma 2. *For any $x > 0$,*

$$\sqrt{x+1/4} < \frac{\Gamma(x+1)}{\Gamma(x+1/2)} < \sqrt{x+1/2}. \tag{7}$$

Proof. For an integer x, (7) was proved in Lorch (1984) and for an arbitrary x in Laforgia (1984). □

Many useful inequalities for the gamma function are in a recent paper Laforgia and Natalini (2011). We shall need (7) for semi-integer x.

2 Mean Length of Confidence Intervals

2.1 Confidence interval for the scale parameter of gamma distribution

Let now $(X_1, X_2, \ldots, X_n)$ be a sample from a population $Gamma\,(a, \beta)$ with a known shape parameter a and a scale parameter β to be estimated. The sum $\sum_{i=1}^{n} X_i \sim Gamma\,(na, \beta)$ is a sufficient statistics for β and the ratio $\sum_{i=1}^{n} X_i/\beta \sim Gamma\,(na, 1)$ is a pivot leading to the standard confidence interval for β of level $1-\alpha$, $\left(\frac{n\overline{X}_n}{\gamma_{na;\,1-\alpha/2}}, \ \frac{n\overline{X}_n}{\gamma_{na;\,\alpha/2}}\right)$. Its average length is $L_n = \beta\left(\frac{1}{\gamma_{na;\,\alpha/2}/na} - \frac{1}{\gamma_{na;\,1-\alpha/2}/na}\right)$, where $\gamma_{na;\,\alpha/2}$ is the quantile of order $\alpha/2$ of $Gamma\,(na, 1)$.

If $G_i(x)$ is the distribution function of $X_i \sim Gamma(a_i, 1)$, $i = 1, 2$, then for $a_1 < a_2$, $G_1(x) > G_2(x)$ (equivalently, X_1 is stochastically smaller than X_2). Therefore, $\gamma_{a_1;\,\alpha} < \gamma_{a_2;\,\alpha}$ for all α, $0 < \alpha < 1$. In particular,

$$\gamma_{na;\,\alpha} < \gamma_{(n+1)a;\,\alpha}. \tag{8}$$

The quantile of order α of $Gamma\,(na, 1/na)$ is $\gamma_{na;\,\alpha}/na$ and its relation to the quantile $\gamma_{(n+1)a;\,\alpha}/(n+1)a$ differs from (8). The following result holds.

Theorem 1. *For $\alpha < \alpha^*(n, a)$, $L_{n+1} < L_n$.*

Proof. By virtue of Corollary 1 applied to the case of $a_1 = na, \beta_1 = 1/na, a_2 = (n+1)a, \beta_2 = 1/(n+1)a$ one gets

$$\frac{1}{\gamma_{(n+1)a;\,\alpha/2}/(n+1)a} < \frac{1}{\gamma_{na;\,\alpha/2}/na} \quad \text{and} \quad \frac{1}{\gamma_{(n+1)a;\,1-\alpha/2}/(n+1)a} > \frac{1}{\gamma_{na;\,1-\alpha/2}/na}, \tag{9}$$

for $\alpha < \alpha^*(n, a)$. The result follows immediately from (9). □

As a function of n for a given a, $\alpha^*(n, a)$ grows very fast (see Figure 1). An interesting fact is that if the sample means $\overline{X}_n$, and $\overline{X}_{n+1}$ calculated from samples of size n and $n+1$,

are the same, the confidence interval $\left(\frac{(n+1)\overline{X}_{n+1}}{\gamma_{(n+1)a;\,1-\alpha/2}}, \; \frac{(n+1)\overline{X}_{n+1}}{\gamma_{(n+1)a;\,\alpha/2}}\right)$ of level $1-\alpha$ for the scale parameter β, lies inside the interval $\left(\frac{n\overline{X}_n}{\gamma_{na;\,1-\alpha/2}}, \; \frac{n\overline{X}_n}{\gamma_{na;\,\alpha/2}}\right)$.

Note that Theorem 1 also holds for an asymmetric confidence interval. Namely, let $\alpha_1 + \alpha_2 = \alpha$, then the average length of the confidence interval $\left(\frac{n\overline{X}_n}{\gamma_{na;\,1-\alpha_2}}, \; \frac{n\overline{X}_n}{\gamma_{na;\,\alpha_1}}\right)$ is a decreasing function of n.

A standard (one-sided) lower confidence bound of level $1-\alpha$ for the parameter β is $\frac{n\overline{X}_n}{\gamma_{na;\,1-\alpha}}$. The statistician is interested in having (for a given level $1-\alpha$) a larger lower bound. From Corollary 1 for $\alpha < \alpha*$ it follows that $E\left(\frac{n\overline{X}_n}{\gamma_{na;\,1-\alpha}}\right) = \frac{\beta}{\gamma_{na;\,1-\alpha}/na}$ is an increasing function of n.

2.2 Confidence Interval for the Normal Variance

Let now $(X_1, X_2, \ldots, X_n)$ be a sample from a normal population $N(\mu, \sigma^2)$ with μ and σ^2 as parameters. The standard confidence interval of level $(1-\alpha)$ for σ^2 is

$$\left(\frac{(n-1)S_n^2}{\chi^2_{n-1;\,1-\alpha/2}}, \; \frac{(n-1)S_n^2}{\chi^2_{n-1;\,\alpha/2}}\right), \tag{10}$$

where S_n^2 is the sample variance and $\chi^2_{n-1;\,\alpha/2}$ is the quantile of order $\alpha/2$ of chi-square distribution with $n-1$ degrees of freedom. The average length of the interval (10) is $L_n = \sigma^2\left(\frac{1}{\chi^2_{n-1;\,\alpha/2}/(n-1)} - \frac{1}{\chi^2_{n-1;\,1-\alpha/2}/(n-1)}\right)$. If $X \sim \chi^2_d$, then $X/d \sim Gamma\left(\frac{d}{2}, \frac{2}{d}\right)$ and again Corollary 1 is applicable. Thus, for $\alpha < \alpha^*(n)$, monotonicity of L_n holds, $L_n > L_{n+1}$. A table of the values of $\alpha^*(n)$ can be found in Székely and Bakirov (2003). For the sake of completeness a graph of $\alpha^*(n)$ is drawn in Figure 2.2.

2.3 Confidence Interval for the Normal Mean

The standard Student confidence interval of level $1-\alpha$ for μ is

$$\left(\overline{X}_n - t_{n-1;\,\alpha/2}\frac{S_n}{\sqrt{n}}, \; \overline{X}_n + t_{n-1;\,\alpha/2}\frac{S_n}{\sqrt{n}}\right), \tag{11}$$

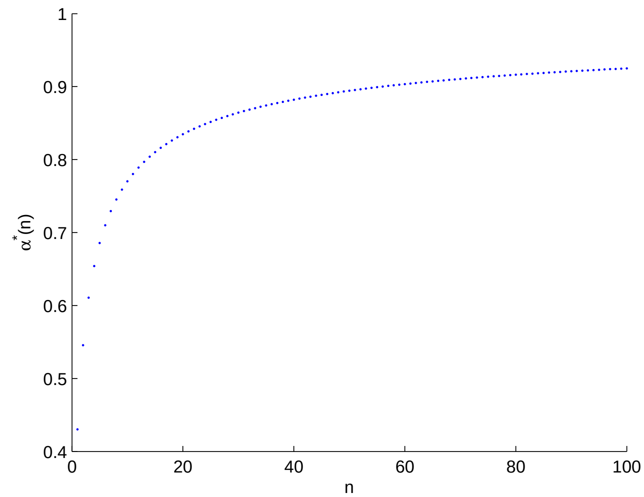

Figure 2: The critical values of α^* as a function of n

where $t_{d;\,\alpha}$ is the quantile of order $1-\alpha$ of the Student distribution with d degrees of freedom. The average length L_n of (11) is easily calculated,

$$L_n = 2\sqrt{2}\sigma \;\; t_{n-1;\,\alpha/2} E_n, \;\; \text{where} \;\; E_n = \frac{\Gamma(n/2)}{\Gamma((n-1)/2)\sqrt{n(n-1)}}. \tag{12}$$

The quantile $t_{d;\,\alpha}$ monotonically decrease in d. This presumable known fact can be easily proved along the lines of Lemma 1. Indeed, let $Z \sim N(0,1)$ and $\xi_n \sim \chi^2_n/n$ be independent random variables. Then for any $c > 0$,

$$P\left(\frac{Z}{\sqrt{\xi_n}} > c\right) = E\left\{P\left(\frac{Z}{\sqrt{\xi_n}} > c \,|\, Z\right)\right\} = E\left\{P\left(\xi_n < Z^2/c^2 \,|\, Z\right)\right\} > E\left\{P\left(\xi_{n+1} < Z^2/c^2 \,|\, Z\right)\right\}$$
$$= E\left\{P\left(\frac{Z}{\sqrt{\xi_{n+1}}} > c \,|\, Z\right)\right\} = P\left(\frac{Z}{\sqrt{\xi_{n+1}}} > c\right),$$

where independence of Z and ξ_n (ξ_{n+1}) and Corollary 1 are used. To prove that $E_n > E_{n+1}$ take first the left inequality from Lemma 2. One has

$$E_n = \frac{1}{\sqrt{n(n-1)}}\frac{\Gamma(n/2)}{\Gamma((n-1)/2)} = \frac{1}{\sqrt{n(n-1)}}\frac{\Gamma((n-2)/2+1)}{\Gamma((n-2)/2+1/2)} > \frac{1}{\sqrt{n(n-1)}}\sqrt{\frac{n-2}{2}+\frac{1}{4}}. \tag{13}$$

The right inequality from Lemma 2 implies

$$E_{n+1} = \frac{1}{\sqrt{n(n+1)}} \frac{\Gamma((n+1)/2)}{\Gamma(n/2)} = \frac{1}{\sqrt{n(n+1)}} \frac{\Gamma((n-1)/2+1)}{\Gamma((n-1)/2+1/2)} < \frac{1}{\sqrt{n(n+1)}} \sqrt{\frac{n-1}{2}+\frac{1}{2}}. \tag{14}$$

Now comparing the right hand sides of (13) and (14) results in $E_n > E_{n+1}$ for $n > 3$. For $n = 2, 3$, the inequalities $E_2 > E_3 > E_4$ follow from the explicitly calculated values of E_2, E_3, and E_4.

References

Artstein, S., Ball, K. M., Barthe, F., Naor, A. (2004). Solution of Shannon's problem on the monotonicity of entropy. *J. Amer. Math. Soc.,* **17,** 975–982.

DasGupta, A. (2008). Letter to the Editors. *IMS Bulletin,* **36,** No. **6,** 16.

Kagan, A. M., Yu, T. (2009). A geometric property of the sample mean and residuals. *Stat. Probab. Letters,* **79,** 1409–1413.

Kagan, A. M., Yu, T., Barron, A., Madiman, M. (2011). Contribution to the theory of Pitman estimators. *Preprint.*

Laforgia, A. (1984). Further inequalities for the gamma function. *Math. Comp.,* **42,** 597–600.

Laforgia, A., Natalini, P. (2011). Some inequalities for the ratio of gamma functions. *J. Ineq. Spec. Fun.,* **2,** 16–26.

Lorch, L. (1984). Inequalities for ultraspherical polynomials and the gamma function. *J. Approx. Theory,* **40,** 115–120.

Shi, N-Z. (2008). Letter to the Editors. *IMS Bulletin,* **36,** No. **4,** 4.

Székely, G. J., Bakirov, N. K. (2003). Extremal probabilities for Gaussian quadratic forms. *Probab. Theory Relat. Fields,* **126,** 184–202.